

# Remarks on multi-marginal symmetric Monge-Kantorovich problems

Nassif Ghoussoub\*

*Department of Mathematics  
University of British Columbia  
Vancouver BC Canada V6T 1Z2  
nassif@math.ubc.ca*

Bernard Maurey†

*Institut de Mathématiques, UMR 7586 - CNRS  
Université Paris Diderot - Paris 7  
Paris, France  
maurey@math.jussieu.fr*

December 10, 2012

## Abstract

Symmetric Monge-Kantorovich transport problems involving a cost function given by a family of vector fields were used by Ghoussoub-Moameni to establish polar decompositions of such vector fields into  $m$ -cyclically monotone maps composed with measure preserving  $m$ -involutions ( $m \geq 2$ ). In this note, we relate these symmetric transport problems to the Brenier solutions of the Monge and Monge-Kantorovich problem, as well as to the Gangbo-Świąch solutions of their multi-marginal counterparts, both of which involving quadratic cost functions.

## 1 Introduction

Given Borel probability measures  $\mu_i$ ,  $i = 0, 1, \dots, m-1$  on domains  $\Omega_i \subset \mathbb{R}^d$ , and a cost function  $c : \Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1} \rightarrow \mathbb{R}$ , the multi-marginal version of Monge's optimal transportation problem is to minimize:

$$C(T_1, \dots, T_{m-1}) := \int_{\Omega_0} c(x_0, T_1(x_0), T_2(x_0), \dots, T_{m-1}(x_0)) d\mu_0 \quad (\mathbf{M})$$

among all  $(m-1)$ -tuples of measurable maps  $(T_1, T_2, \dots, T_{m-1})$ , where  $T_i : \Omega_0 \rightarrow \Omega_i$  pushes  $\mu_0$  forward to  $\mu_i$  for all  $i = 1, \dots, m-1$ . The Kantorovich formulation of the problem is to minimize:

$$C(\theta) := \int_{\Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1}} c(x_0, x_1, x_2, \dots, x_{m-1}) d\theta \quad (\mathbf{K})$$

among all probability measures  $\theta$  on  $\Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1}$  such that the canonical projection

$$\pi_i : \Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1} \rightarrow \Omega_i$$

pushes  $\theta$  forward to  $\mu_i$  for all  $i$ .

Note that for any  $(m-1)$ -tuple  $(T_1, T_2, \dots, T_{m-1})$  such that  $T_i \# \mu_0 = \mu_i$  for all  $i = 1, 2, \dots, m-1$ , we can define the measure  $\theta = (I, T_1, T_2, \dots, T_{m-1}) \# \mu_0$  on  $\Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1}$ , where  $I : \Omega_0 \rightarrow \Omega_0$  is the identity map. Then  $\theta$  projects to  $\mu_i$  for all  $i$  and  $C(T_1, T_2, \dots, T_{m-1}) = C(\theta)$ . In other words,  $(\mathbf{K})$  can be interpreted as a relaxed version of  $(\mathbf{M})$ .

Standard results for fairly general cost functions  $c$  show that there exists a probability measure  $\bar{\theta}$  on  $\Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1}$  with marginals  $\mu_i$ ,  $i = 0, 1, \dots, m-1$ , where the supremum in  $(\mathbf{K})$  is attained. The natural question here is the following:

---

\*Partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

†Research supported by CNAV93.

**Problem (1):** For which cost functions  $c$ , problem **(K)** admits a solution  $\bar{\theta}$  that is supported on a “graph”, that is a measure of the form  $\bar{\theta} = (I, T_1, T_2, \dots, T_{m-1})_{\#} \mu_0$  for a suitable family of point transformations  $(T_1, T_2, \dots, T_{m-1})$ .

Whereas the case when  $m = 2$  is already well understood, the problem when  $m \geq 3$  remains elusive since existence and uniqueness in **(M)** as well as uniqueness in **(K)** are still largely open for general cost functions. There is however one important case where this problem has been resolved by Gangbo and Świąch [7]. This is when the cost function is given by

$$c(x_0, x_1, \dots, x_{m-1}) = \sum_{i=0}^{m-1} \sum_{j=i+1}^{m-1} |x_i - x_j|^2. \quad (1)$$

Other extensions were also given by Pass [16, 17, 18, 19] and by Carlier-Nazaret [4].

In this note, we are interested in the symmetric versions of the Monge-Kantorovich problem. They are of the following type.

$$\mathbf{K}_{\text{sym}} = \sup \left\{ \int_{\Omega^m} c(x_0, x_1, \dots, x_{m-1}) d\theta; \theta \in \mathcal{P}_{\text{sym}}(\Omega^m, \mu) \right\} \quad (2)$$

where  $c$  is an appropriate cost function and  $\mathcal{P}_{\text{sym}}(\Omega^m, \mu)$  denotes the set of all probability measures on  $\Omega^m$ , which are invariant under the cyclical permutation

$$\sigma(x_0, x_1, \dots, x_{m-1}) = (x_1, x_2, \dots, x_{m-1}, x_0)$$

and whose marginals are all equal to a given measure  $\mu$ . Note that one can then assume that the cost function  $c$  is cyclically symmetric, since one can replace it by its symmetrization

$$\tilde{c}(\mathbf{x}) = \frac{1}{m} \sum_{i=0}^{m-1} c(\sigma^i(\mathbf{x})), \quad (3)$$

and in this case, one can minimize over the set  $\mathcal{P}(\Omega^m, \mu)$  of all probability measures on  $\Omega^m$  whose all marginals are equal to  $\mu$ .

Standard results for fairly general cost functions  $c$  show that there exists  $\bar{\theta} \in \mathcal{P}_{\text{sym}}(\Omega^m, \mu)$ , where the supremum above is attained. The natural question here is the following:

**Problem (2):** For which cost functions  $c$ , problem **(K<sub>sym</sub>)** admits as a solution a probability measure  $\bar{\theta}$  of the form  $\bar{\theta} = (I, S, S^2, \dots, S^{m-1})_{\#} \mu$ , where  $S$  is a  $\mu$ -measure preserving transformation on  $\Omega$  such that  $S^m = I$  a.e.

Problem (2) was resolved by Ghoussoub and Moameni for  $m = 2$  in [13] and for  $m \geq 3$  in [14] in the case where the cost function is of the form

$$c(x_0, x_1, \dots, x_{m-1}) = \langle u_1(x_0), x_1 \rangle + \dots + \langle u_{m-1}(x_0), x_{m-1} \rangle, \quad (4)$$

where  $u_1, \dots, u_{m-1}$  are bounded vector fields from  $\Omega \rightarrow \mathbb{R}^d$ . Their work was in the context of establishing polar decompositions of vector fields in terms of monotone operators, which we will briefly describe below. The raison-d'être of this paper is however to make a link between the results of Gangbo and Świąch dealing with the quadratic cost (1) but for marginals of the form  $\mu_i = \sigma_{\#}^i \mu$  for  $i = 0, \dots, m-1$ , and the symmetric Monge-Kantorovich problems considered by Ghoussoub and Moameni for the cost (4).

## Polar decompositions

Recall that a vector field  $u : \Omega \rightarrow \mathbb{R}^d$  on a domain  $\Omega$  in  $\mathbb{R}^d$  is said to be monotone on  $\Omega$  if for all  $(x, y)$  in  $\Omega$ ,

$$\langle x - y, u(x) - u(y) \rangle \geq 0. \quad (5)$$

A result of E. Krauss [15] states that a map  $u : \Omega \rightarrow \mathbb{R}^d$  is *monotone* if and only if

$$u(x) = \nabla_2 H(x, x) \text{ for all } x \in \Omega, \quad (6)$$

where  $H$  is a concave-convex anti-symmetric Hamiltonian on  $\mathbb{R}^d \times \mathbb{R}^d$ . More recently, Galichon-Ghoussoub [6] extended Krauss' result to the case of *m-cyclically monotone vector fields*, where  $m$  is a fixed integer larger than 2. Recall that these are the maps  $u$  from  $\Omega$  to  $\mathbb{R}^d$  that satisfy for any  $m+1$  points  $(x_i)_{i=0}^m$  in  $\Omega$  with  $x_0 = x_m$ , the inequality

$$\sum_{k=0}^{m-1} \langle u(x_{k+1}), x_{k+1} - x_k \rangle \geq 0. \quad (7)$$

For that, Galichon and Ghoussoub consider the class  $\mathcal{H}_m(\Omega)$  of all *m-cyclically antisymmetric Hamiltonians* on  $\Omega^m$ , that is the set

$$\mathcal{H}_m(\Omega) = \{H \in C(\Omega^m; \mathbb{R}); \sum_{i=1}^m H(\sigma^{i-1}(\mathbf{x})) = 0 \text{ for all } \mathbf{x} \in \Omega^m\},$$

where  $\sigma$  is the cyclical permutation  $\sigma(x_0, x_1, \dots, x_{m-1}) = (x_1, x_2, \dots, x_{m-1}, x_0)$ . They then show that if a vector field  $u$  is *m-cyclically monotone*, then there exists a Hamiltonian  $H \in \mathcal{H}_m(\Omega)$  such that

$$u(x) = \nabla_m H(x, x, \dots, x) \text{ for all } x \in \Omega. \quad (8)$$

Moreover,  $H$  can be assumed to be concave in the first variable, convex in the last  $(m-1)$  variables, though only *m-cyclically sub-antisymmetric* on  $\Omega^m$ , that is  $\sum_{i=1}^m H(\sigma^{i-1}(\mathbf{x})) \leq 0$  for all  $\mathbf{x} \in \Omega^m$ .

It is worth comparing the above to a classical theorem of Rockafellar [20], which yields that a single-valued map  $u$  from  $\Omega$  to  $\mathbb{R}^d$  is a *maximal cyclically monotone operator* (i.e., satisfies (7) for every  $m \geq 2$ ), if and only if

$$u(x) = \nabla \varphi(x) \text{ on } \Omega, \text{ where } \varphi : \mathbb{R}^d \rightarrow \mathbb{R} \text{ is a convex function.} \quad (9)$$

More remarkable is the polar decomposition that Y. Brenier [3] establishes for a general non-degenerate vector field, and which follows from his celebrated mass transport theorem. Recall that a mapping  $u : \Omega \rightarrow \mathbb{R}^d$  is said to be *non-degenerate* if the inverse image  $u^{-1}(N)$  of every zero-measure  $N \subseteq \mathbb{R}^d$  has also zero measure. Brenier proved that any non-degenerate vector field  $u \in L^\infty(\Omega, \mathbb{R}^d)$  can be decomposed as

$$u(x) = \nabla \varphi \circ S(x) \quad \text{a.e. in } \Omega, \quad (10)$$

with  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  being a convex function and  $S : \bar{\Omega} \rightarrow \bar{\Omega}$  a measure preserving transformation.

Later, and in the same spirit as Brenier's, Ghoussoub and Moameni established in [13] another decomposition for non-degenerate vector fields, which can be seen as the general version of Krauss' characterization of monotone operators. Assuming the boundary  $\partial\Omega$  has measure zero, they show that if  $u \in L^\infty(\Omega, \mathbb{R}^d)$  is a non-degenerate vector field, then there exists a measure preserving transformation  $S : \Omega \rightarrow \Omega$  such that  $S^2 = I$  (i.e., an involution), and a globally Lipschitz anti-symmetric concave-convex Hamiltonian  $H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that

$$u(x) = \nabla_2 H(x, Sx) \quad \text{a.e. } x \in \Omega. \quad (11)$$

In other words, up to a measure preserving involution, essentially every bounded vector field is monotone, where the latter correspond to when  $S$  is the identity map on  $\Omega$ .

More recently, Ghoussoub-Moameni [14] extended this result by showing that any family  $u_1, \dots, u_{m-1}$  of non-degenerate bounded vector field can be represented as

$$u_i(x) = \nabla_{i+1} H(x, Sx, S^2x, \dots, S^{m-1}x) \quad \text{a.e. } x \in \Omega \quad \text{for } i = 1, \dots, m-1, \quad (12)$$

where  $H \in \mathcal{H}_m(\Omega)$  and  $S$  is a measure preserving *m*-involution (i.e.,  $S^m = I$ ). Moreover,  $H$  could be replaced by a Hamiltonian that is concave in the first variable and convex in the other  $(m-1)$ -variables, though only *m-cyclically sub-antisymmetric*.

The proofs of the representations (11) (when  $m = 2$ ) and of (12) (when  $m \geq 3$ ) rely on symmetric versions of the Monge problem and of its multi-marginal Monge-Kantorovich version as mentioned above. We shall give here another formulation that is closer to the original Monge and multi-marginal Monge-Kantorovich problems corresponding to a quadratic cost.

## 2 Mass transport with quadratic cost in the presence of symmetry

Given two probability measures with finite second moment  $\mu_0, \mu_1$  on  $\mathbb{R}^d$ , with  $X := \text{support}(\mu_0)$  and  $Y := \text{support}(\mu_1)$ , the Wasserstein distance  $W_2(\mu_0, \mu_1)$  between them is defined by the formula

$$W_2(\mu_0, \mu_1)^2 = \inf \left\{ \int_X |x - T(x)|^2 d\mu_0(x); T \in \mathcal{S}(\mu_0, \mu_1) \right\} \quad (13)$$

where  $\mathcal{S}(\mu_0, \mu_1)$  is the class of all Borel measurable maps  $T : X \rightarrow Y$  such that  $T_{\#}\mu_0 = \mu_1$ , i.e., those which satisfy the change of variables formula,

$$\int_Y h(y) d\mu_1(y) = \int_X h(T(x)) d\mu_0(x), \quad \text{for every } h \in C(Y). \quad (14)$$

Whether the infimum describing the Wasserstein distance  $W_2(\mu_0, \mu_1)$  is achieved by an optimal map  $\bar{T}$  is a variation on the original mass transport problem of G. Monge, who inquired about finding the optimal way for rearranging  $\mu_0$  into  $\mu_1$  against the cost function  $c(x) = |x|$ . Our cost function here  $c(x) = \frac{1}{2}|x|^2$  is quadratic, and the existence, uniqueness and characterization of an optimal map that we give below, was established by Y. Brenier.

**Theorem 2.1 (Brenier)** *Assume  $\mu_0$  is absolutely continuous with respect to Lebesgue measure, then there exists a unique optimal map  $\bar{T}$  in  $\mathcal{S}(\mu_0, \mu_1)$ , where the infimum in (13) is achieved. Moreover, the map  $\bar{T} : X \rightarrow Y$  is one-to-one and onto,  $\mu_0$  a.e., and is equal to  $\nabla\varphi$   $\mu_0$  a.e on  $X$ , for some convex function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ .*

*Moreover, the Brenier map  $\bar{T}$  is the unique map (up to  $\mu_0$  a.e. equivalence) of the form  $\nabla\varphi$  with  $\varphi$  convex such that  $\nabla\varphi_{\#}\mu_0 = \mu_1$ .*

Now we consider the above theorem in the presence of symmetry.

**Corollary 2.2** *Let  $\mu$  be a probability measure on  $\mathbb{R}^d$  that is absolutely continuous with respect to Lebesgue measure, and let  $\tilde{\mu}$  be its image by a self-adjoint unitary transformation  $\sigma$  on  $\mathbb{R}^d$  (i.e.,  $\sigma = \sigma^*$  and  $\sigma^2 = I$ ). Then, there exists a convex function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  such that*

$$\nabla\varphi_{\#}\mu = \tilde{\mu} \quad \text{and} \quad \varphi^*(\sigma(x)) = \varphi(x) \quad \text{for } x \in \mathbb{R}^d, \quad (15)$$

where  $\varphi^*$  is the Legendre transform of  $\varphi$ .

**Proof:** The above theorem yields a convex function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $\nabla\varphi_{\#}\mu = \tilde{\mu}$ . Recall that if  $\varphi^*$  is the Legendre transform of  $\varphi$ , then  $\nabla\varphi^* = (\nabla\varphi)^{-1}$ . Hence the function  $\psi(x) = \varphi^*(\sigma(x))$ , which is also convex has a gradient  $\nabla\psi = \sigma^* \circ \nabla\varphi^* \circ \sigma$ , which also maps  $\mu$  onto  $\tilde{\mu}$ . By the uniqueness property, we have  $\nabla\psi = \nabla\varphi$ , which means that –up to a constant–  $\varphi^*(\sigma(x)) = \varphi(x)$  for all  $x \in \mathbb{R}^d$ .

In the rest of this section, we try to connect the above corollary to the following polar decomposition established in [13].

**Theorem 2.3 (Ghoussoub-Moameni)** *Let  $\Omega$  be an open bounded set in  $\mathbb{R}^d$  such that  $\partial\Omega$  has zero Lebesgue measure.*

1. *If  $u \in L^\infty(\Omega, \mathbb{R}^d)$  is a non-degenerate vector field, then there exists a measure preserving transformation  $S : \bar{\Omega} \rightarrow \bar{\Omega}$  such that  $S^2 = I$  (i.e., an involution), and a globally Lipschitz anti-symmetric concave-convex Hamiltonian  $H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that*

$$u(x) = \nabla_2 H(x, Sx) \quad \text{a.e. } x \in \Omega. \quad (16)$$

*The involution  $S$  is obtained by solving the following variational problem*

$$\sup \left\{ \int_{\Omega} \langle u(x), Sx \rangle dx; S \text{ is a measure preserving involution on } \Omega \right\}. \quad (17)$$

2. *Moreover,  $u$  is a monotone map if and only if there is a representation as (16) with  $S$  being the identity.*

## Lagrangians and Hamiltonians associated to monotone maps

An important example of an involutive transformation is the transpose  $\sigma$  on  $\mathbb{R}^d \times \mathbb{R}^d$  defined by  $\sigma(x, y) = (y, x)$ , since the convex functions  $L$  on  $\mathbb{R}^d \times \mathbb{R}^d$  satisfying  $L^*(y, x) = L(x, y)$  for all  $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$  are connected to central notions in nonlinear analysis and PDEs [11].

Duality theory is at the heart of this concept and it is therefore enlightening to describe it in the case where  $\mathbb{R}^d$  is replaced by any reflexive Banach space  $X$ . Recall from [11] the notion of a vector field  $\bar{\partial}L$  that is derived from a convex lower semi-continuous Lagrangian on phase space  $L : X \times X^* \rightarrow \mathbb{R} \cup \{+\infty\}$  in the following way: for each  $x \in X$ , the –possibly empty– subset  $\bar{\partial}L(x)$  of  $X^*$  is defined as

$$\bar{\partial}L(x) := \{p \in X^*; (p, x) \in \partial L(x, p)\}. \quad (18)$$

Here  $\partial L$  is the subdifferential of the convex function  $L$  on  $X \times X^*$ , which should not be confused with  $\bar{\partial}L$ . Of particular interest are those vector fields derived from *self-dual Lagrangians*, i.e., those convex lower semi-continuous Lagrangians  $L$  on  $X \times X^*$  that satisfy the following duality property:

$$L^*(p, x) = L(x, p) \quad \text{for all } (x, p) \in X \times X^*, \quad (19)$$

where here  $L^*$  is the Legendre transform in both variables, i.e.,

$$L^*(p, x) = \sup\{\langle y, p \rangle + \langle x, q \rangle - L(y, q) : (y, q) \in X \times X^*\}.$$

Such Lagrangians satisfy the following basic property:

$$L(x, p) - \langle x, p \rangle \geq 0 \quad \text{for every } (x, p) \in X \times X^*. \quad (20)$$

Moreover,

$$L(x, p) - \langle x, p \rangle = 0 \quad \text{if and only if } (p, x) \in \partial L(x, p), \quad (21)$$

which means that the associated vector field at  $x \in X$  is simply

$$\bar{\partial}L(x) := \{p \in X^*; L(x, p) - \langle x, p \rangle = 0\}. \quad (22)$$

These so-called *selfdual vector fields* are natural but far reaching extensions of subdifferentials of convex lower semi-continuous functions. Indeed, the most basic selfdual Lagrangians are of the form

$$L(x, p) = \varphi(x) + \varphi^*(p),$$

where  $\varphi$  is a convex and lower semi-continuous function on  $X$ , and  $\varphi^*$  is its Legendre conjugate on  $X^*$ , in which case

$$\bar{\partial}L(x) = \partial\varphi(x).$$

More interesting examples of self-dual Lagrangians are of the form

$$L(x, p) = \varphi(x) + \varphi^*(-\Gamma x + p),$$

where  $\varphi$  is as above, and  $\Gamma : X \rightarrow X^*$  is a skew adjoint operator. The corresponding selfdual vector field is then

$$\bar{\partial}L(x) = \Gamma x + \partial\varphi(x). \quad (23)$$

Actually, it turned out that any *maximal monotone operator*  $A$  is a self-dual vector field and vice-versa [11]. That is, there exists a selfdual Lagrangian  $L$  such that  $A = \bar{\partial}L$ . This fact was proved and reproved by several authors. See for example, R. S. Burachik and B. F. Svaiter [2], B. F. Svaiter [22].

This result means that self-dual Lagrangians can be seen as the *potentials* of maximal monotone operators, in the same way as the Dirichlet integral is the potential of the Laplacian operator (and more generally as any convex lower semi-continuous energy is a potential for its own subdifferential). Check [11] to see how this characterization leads to variational formulations and resolutions of most equations involving monotone operators.

**Proposition 2.1** *Let  $u : \Omega \rightarrow \mathbf{R}^d$  be a possibly set-valued map. The following properties are then equivalent:*

1.  $u$  is a maximal monotone map with domain  $\Omega$ .

2. There exists a convex self-dual Lagrangian on  $\mathbb{R}^d \times \mathbb{R}^d$  such that  $u(x) = \bar{\partial}L(x)$  for all  $x \in \Omega$ . In other words,

$$(p, x) \in \partial L(x, p) \text{ if and only if } p \in u(x). \quad (24)$$

3. There exists a concave-convex anti-symmetric Hamiltonian  $H$  on  $\mathbb{R}^d \times \mathbb{R}^d$  such that

$$u(x) \subset \partial_2 H(x, x) \text{ for } x \in \Omega. \quad (25)$$

**Sketch of proof:** Assuming  $u$  is maximal monotone with domain  $\Omega$ , we consider its associated Fitzpatrick function, that is

$$N(p, x) = \sup\{\langle p, y \rangle + \langle q, x - y \rangle; (y, q) \in \text{Graph}(u)\}. \quad (26)$$

It is known and easy to see that

$$N^*(x, p) \geq N(p, x) \geq \langle x, p \rangle \text{ for every } (x, p) \in \mathbb{R}^d \times \mathbb{R}^d, \quad (27)$$

and that

$$N(p, x) = \langle x, p \rangle \text{ if and only if } (x, p) \in \text{Graph}(u).$$

Now consider the following Lagrangian on  $\mathbb{R}^d \times \mathbb{R}^d$ , which interpolates between  $N$  and  $N^*$ ,

$$L(p, x) := \inf \left\{ \frac{1}{2}N(p_1, x_1) + \frac{1}{2}N^*(x_2, p_2) + \frac{1}{8}\|x_1 - x_2\|^2 + \frac{1}{8}\|p_1 - p_2\|^2 \right\},$$

where the infimum is taken over all couples  $(x_1, p_1)$  and  $(x_2, p_2)$  such that

$$(x, p) = \frac{1}{2}(x_1, p_1) + \frac{1}{2}(x_2, p_2).$$

It can be shown (see [11], p. 92) that  $L$  is a self-dual Lagrangian on  $\mathbb{R}^d \times \mathbb{R}^d$ , in such a way that

$$N^*(x, p) \geq L(p, x) \geq N(p, x) \geq \langle x, p \rangle \text{ for every } (x, p) \in \mathbb{R}^d \times \mathbb{R}^d, \quad (28)$$

which means that if  $(x, p) \in \Omega \times \mathbb{R}^d$ , then  $L(p, x) = \langle x, p \rangle$  if and only if  $(x, p) \in \text{Graph}(u)$ , that is when  $p \in u(x)$ .

To show that 2) implies 3), it suffices to take the Legendre transform of  $L$  with respect to the second variable, i.e.,

$$K_L(x, y) = \sup\{\langle y, p \rangle - L(x, p); p \in \mathbb{R}^d\}.$$

It is clearly concave-convex. The selfduality of  $L$  yields that  $K_L$  is (at least) sub-antisymmetric, i.e.,

$$K_L(x, y) \leq -K_L(y, x) \text{ for } (x, y) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Note that  $K_L(x, y) \leq H_L(x, y) = \frac{1}{2}(K_L(x, y) - K_L(y, x))$  is anti-symmetric, and that  $u(x) \in \partial_2 H_L(x, x)$  in  $\Omega$  as in the theorem of Krauss mentioned above.

Finally, assuming 3) that is  $u(x) \in \partial_2 H(x, x)$ , where  $H$  is convex in the second variable, we have for any  $x, y \in \Omega$ , any  $p \in \partial_2 H_L(x, x)$  and  $q \in \partial_2 H_L(y, y)$

$$H(x, y) \geq H(x, x) + \langle p, y - x \rangle \text{ and } H(y, x) \geq H(y, y) + \langle q, x - y \rangle.$$

Since  $H$  is anti-symmetric (hence  $H(x, x) = 0$ ), this implies –by adding the two inequalities– that

$$0 \geq \langle u(x) - u(y), y - x \rangle,$$

which means that  $u$  is monotone. □

Now we note that Monge transport problems provide a natural way to construct selfdual Lagrangians, hence general monotone operators.

**Corollary 2.4** *Let  $\mu$  be a probability measure on the product space  $\Omega := \Omega_1 \times \Omega_2 \subset \mathbb{R}^d \times \mathbb{R}^d$ , and let  $\tilde{\mu}$  be the probability measure on  $\tilde{\Omega} := \Omega_2 \times \Omega_1$  obtained as the image of  $\mu$  by the transformation  $\sigma(x, y) = (y, x)$ . Then, there exists a self-dual Lagrangian  $L : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $\nabla L_{\#}\mu = \tilde{\mu}$ . Moreover, the Monge-Kantorovich problem*

$$\sup\left\{\int_{\Omega \times \tilde{\Omega}} [\langle x, y' \rangle + \langle y, x' \rangle] d\theta(x, y, y', x'); \theta \in \mathcal{P}(\Omega \times \tilde{\Omega}), \text{proj}_1\theta = \mu, \text{proj}_2\theta = \tilde{\mu}\right\} \quad (29)$$

*has a solution  $\theta$  that is supported on the self-dual Lagrangian manifold*

$$\{((x, y), (x', y')) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}; L(x, y) + L(x', y') = \langle x, y' \rangle + \langle y, x' \rangle\}. \quad (30)$$

**Proof:** This is a direct application of the above corollary. The map  $\nabla L$  then solves the Monge problem

$$\int_{\Omega_1 \times \Omega_2} |(x, y) - \nabla L(x, y)|^2 d\mu(x, y) = \inf\left\{\int_{\Omega_1 \times \Omega_2} |(x, y) - T(x, y)|^2 d\mu(x, y); T_{\#}\mu = \tilde{\mu}\right\}. \quad (31)$$

The self-duality of  $L$  follows from applying the above corollary to the transformation  $\sigma(x, y) = (y, x)$ .  $\square$

## Involutions as a byproduct of mass transport between graphs

Suppose now that  $\mu = (I \otimes u)_{\#}dx$ , where  $dx$  is normalized Lebesgue measure on a bounded domain  $\Omega$  in  $\mathbb{R}^d$ , and  $u : \Omega \rightarrow \mathbb{R}^d$  is a vector field in such a way that  $\mu$  is supported by the graph  $G := \{(x, u(x)); x \in \Omega\}$  and let  $\tilde{\mu} = \sigma_{\#}\mu$ . Now any map  $T$  pushing  $\mu$  onto  $\tilde{\mu}$  can be parameterized by an application  $S : \Omega \rightarrow \Omega$  via the formula:

$$T : (x, u(x)) \rightarrow (u(Sx), Sx), \quad (32)$$

and the Monge problem between  $\mu$  and  $\tilde{\mu}$  can then be formulated as

$$\inf\left\{\frac{1}{2}\int_{\Omega} [|u(Sx) - x|^2 + |u(x) - S(x)|^2] dx; S \text{ measure preserving transformation on } \Omega\right\}. \quad (33)$$

Assume now that –just like in the non-degenerate case– there exists a self-dual Lagrangian  $L$  such that  $\nabla L_{\#}\mu = \tilde{\mu}$ . This means that there exists  $S : \Omega \rightarrow \Omega$  such that

$$\nabla L(x, u(x)) = (u(Sx), Sx) \text{ for a.e. } x \in \Omega.$$

Since  $\nabla L^* = \sigma^* \circ \nabla L \circ \sigma$  and  $\nabla L^*(u(Sx), Sx) = (x, u(x))$  a.e., we have for a.e.  $x \in \Omega$ ,

$$(x, u(x)) = \nabla L^*(u(Sx), Sx) = \sigma^* \circ \nabla L(Sx, u(Sx)) = \sigma^*(u(S^2x), S^2x) = (S^2x, u(S^2x)).$$

It follows that  $S$  is a measure preserving involution on  $\Omega$ . In other words, problem (33) is equivalent to the problem

$$\sup\left\{\int_{\Omega} \langle u(x), Sx \rangle dx; S \text{ measure preserving involution on } \Omega\right\}, \quad (34)$$

which was used by Ghoussoub-Moameni to establish the polar decomposition (11). Actually, if one now considers the Legendre transform  $H$  of  $L$  with respect to the second variable, then as noted above,  $H$  is sub-anti-symmetric, can be assumed to be anti-symmetric, and satisfies,

$$L(x, u(x)) + H(x, Sx) = \langle u(x), Sx \rangle \text{ for all } x \in \Omega.$$

This then yields the polar decomposition  $u(x) = \nabla_2 H(x, Sx)$  a.e., which can be then seen as a (self-dual) mass transport problem between the measure  $\mu$  supported by the graph of  $u$  and its transpose. Unfortunately, the measure  $\mu$  is too degenerate to fall under the framework where we have uniqueness in Brenier's theorem, hence the need to give a direct proof of the result as in [13].

In order to link the polar decomposition with the symmetric Monge-Kantorovich problem, we note that measurable functions  $S : \Omega \rightarrow \Omega$ , whose graphs  $\{(x, Sx); x \in \Omega\}$  are the support of measures in  $\mathcal{P}_{\text{sym}}^{\mu}(\Omega \times \Omega)$  can be characterized in the following way.

**Lemma 2.5** *Let  $S : \Omega \rightarrow \Omega$  be a measurable map, then the following are equivalent:*

1. *The image of  $\mu$  by the map  $x \rightarrow (x, Sx)$  belongs to  $\mathcal{P}_{\text{sym}}^\mu(\Omega \times \Omega)$ .*
2.  *$S$  is  $\mu$ -measure preserving and  $S^2x = x$   $\mu$ -a.e.*
3.  *$\int_\Omega H(Sx, x) d\mu(x) = 0$  for every Borel measurable, bounded and antisymmetric function  $H$  on  $\Omega \times \Omega$ .*

**Proof.** It is clear that 1) implies 3), while 2) implies 1). We now prove that 2) and 3) are equivalent. Assuming first that  $S$  is measure preserving such that  $S^2 = I$  a.e, then for every anti-symmetric  $H$  in  $L^\infty(\Omega \times \Omega)$ , we have

$$\int_\Omega H(x, S(x)) d\mu(x) = \int_\Omega H(S(x), S^2(x)) d\mu(x) = \int_\Omega H(S(x), x) d\mu(x) = - \int_\Omega H(x, S(x)) d\mu(x),$$

hence  $\int_\Omega H(x, S(x)) d\mu(x) = 0$ .

Conversely, if  $\int_\Omega H(x, S(x)) d\mu(x) = 0$  for every anti-symmetric  $H$ , then it suffices to take  $H(x, y) = f(x) - f(y)$ , where  $f$  is any continuous function on  $\Omega$  to conclude that  $S$  is necessarily  $\mu$ -measure preserving. On the other hand, if one considers the anti-symmetric functional

$$H(x, y) = |S(x) - y| - |S(y) - x|,$$

then  $0 = \int_\Omega H(x, S(x)) d\mu(x) = \int |S^2(x) - x| d\mu(x)$ , which clearly yields that  $S$  is an involution  $\mu$ -almost everywhere.  $\square$

We now give the following variational formulation for monotone operators, which shows that they are in some way orthogonal to involutions. Denote by  $\mathcal{S}_2(\Omega, \mu)$  the set of all  $\mu$ -measure preserving involutions on  $\Omega$ . It is an easy exercise to show that  $\mathcal{S}_2(\Omega, \mu)$  is a closed subset of a sphere of  $L^2(\Omega, \mathbb{R}^d)$ . In order to simplify the exposition, we shall assume that  $d\mu$  is Lebesgue measure  $dx$  normalized to be a probability on the bounded open set  $\Omega$ , and  $\mu$  can and will then be dropped from all notation. We shall also assume that the boundary of  $\Omega$  has measure zero.

**Proposition 2.2** *Let  $u : \Omega \rightarrow \mathbf{R}^d$  be a vector field in  $L^2(\Omega, \mathbf{R}^d)$ . The following properties are then equivalent:*

1.  *$u$  is monotone a.e. on  $\Omega$ , that is there exists a measure zero set  $N$  such that the restriction of  $u$  to  $\Omega \setminus N$  is monotone.*
2.  $\sup \left\{ \int_\Omega \langle u(x), Sx - x \rangle dx; S \in \mathcal{S}_2(\Omega, \mu) \right\} = 0$ .
3. *The projection of  $u$  on  $\mathcal{S}_2(\Omega, \mu)$  is the identity map, that is*

$$\int_\Omega |u(x) - x|^2 dx = \inf \left\{ \int_\Omega |u(x) - S(x)|^2 dx; S \in \mathcal{S}_2(\Omega, \mu) \right\}.$$

4.  $\sup \left\{ \int_{\Omega \times \Omega} \langle u(x), y \rangle d\pi(x, y); \pi \in \mathcal{P}_{\text{sym}}(\Omega \times \Omega, dx) \right\} = \int_\Omega \langle u(x), x \rangle dx$ .

**Proof:** Assume  $u$  is a monotone map and use Proposition 2.1 to write it –modulo an obvious abuse in notation– as  $u(x) = \nabla_2 H(x, x)$  a.e. on  $\Omega$ , where  $H$  is anti-symmetric and convex in the second variable. We can write for any measurable map  $S$ ,

$$H(x, S(x)) \geq H(x, x) + \langle \nabla_2 H(x, x), S(x) - x \rangle.$$

It follows that

$$\int_\Omega \langle u(x), x - S(x) \rangle dx = \int_\Omega \langle \nabla_2 H(x, x), x - S(x) \rangle dx \geq \int_\Omega [H(x, x) - H(x, S(x))] dx.$$



Note that  $\int_{\Omega} H(x, S(x)) dx = 0$  since  $S$  is measure preserving and  $S^2 = I$ . Similarly,  $\int_{\Omega} H(x, x) dx = 0$  and it then follows that  $\int_{\Omega} \langle u(x), x - Sx \rangle dx \geq 0$ . Finally, by taking  $Sx = x$ , one can then see that the supremum in 2) is equal to zero.

Suppose now that 2) holds. In order to show 1) i.e., that  $u$  is monotone, consider a pair  $x_1, x_2$  in  $\Omega$  and  $R$  small enough so that  $B(x_i, R) \subset \Omega$  for  $i \in \{1, 2\}$ . Define the measure preserving involution  $S_R$  via

$$S_R(x) = \begin{cases} x - x_2 + x_1 & \text{if } x \in B(x_2, R) \\ x - x_1 + x_2 & \text{if } x \in B(x_1, R) \\ x & \text{otherwise.} \end{cases}$$

Since  $\int_{\Omega} \langle u(x), x - S_R(x) \rangle dx \geq 0$ , we have

$$\left\langle x_1 - x_2, \int_{B_1} u(x_1 + Ry) dy - \int_{B_1} u(x_2 + Ry) dy \right\rangle \geq 0.$$

Since  $u$  is Lebesgue integrable, almost every point  $x \in \Omega$  is a Lebesgue point, which means that  $u(x) = \lim_{R \rightarrow 0} |B|^{-1} \int_B u(x + Ry) dy$ . This leads to  $\langle x_1 - x_2, u(x_1) - u(x_2) \rangle \geq 0$  for a.e.  $x_1, x_2 \in \Omega$ .

By developing the square, it is clear that property 2) is equivalent to 3), which says that the identity map is the projection of  $u$  on the closed subset  $\mathcal{S}_2(\Omega, \mu)$  of the sphere of  $L^2(\Omega, \mathbf{R}^d)$ , that is  $\text{dist}(u, \mathcal{S}_2(\Omega, \mu)) = \|u - I\|_2$ . In other words,  $\int_{\Omega} |u(x) - x|^2 dx = \inf \{ \int_{\Omega} |u(x) - S(x)|^2 dx; S \in \mathcal{S}_2(\Omega, \mu) \}$ .

For 1) implies 4) assume  $u$  is monotone and observe that for any probability  $\pi$  in  $\mathcal{P}(\Omega \times \Omega)$  with marginals  $dx$ , we have

$$\begin{aligned} \int_{\Omega \times \Omega} \langle u(x), y - x \rangle d\pi(x, y) &= \int_{\Omega \times \Omega} \langle u(x) - u(y), y - x \rangle d\pi + \int_{\Omega \times \Omega} \langle u(y), y - x \rangle d\pi \\ &\leq \int_{\Omega \times \Omega} \langle u(y), y - x \rangle d\pi(x, y). \end{aligned}$$

Since  $\pi$  is symmetric, we have then that  $2 \int_{\Omega \times \Omega} \langle u(x), y - x \rangle d\pi(x, y) \leq 0$ . The fact that the supremum is zero follows from simply taking the probability measure supported on the diagonal of  $\Omega \times \Omega$ .

Finally, note that 4) implies 2) by considering for any  $S \in \mathcal{S}_2(\Omega, \mu)$  the symmetric measure  $d\pi$  on  $\Omega \times \Omega$  that is the image of Lebesgue measure by the map  $x \rightarrow (x, Sx)$ . Note that

$$\int_{\Omega \times \Omega} f(x, y) d\pi(x, y) = \int_{\Omega} f(x, Sx) dx = \int_{\Omega} f(Sx, x) dx = \int_{\Omega \times \Omega} f(x, y) d\pi(y, x).$$

□

Back to the case of a general vector field, the above then shows that the variational problem used in the polar decomposition

$$\sup \left\{ \int_{\Omega} \langle u(x), Sx \rangle dx; S \text{ measure preserving involution on } \Omega \right\}, \quad (35)$$

is nothing but a symmetric Monge-Kantorovich problem

$$\sup \left\{ \int_{\Omega \times \Omega} \langle u(x), y \rangle d\pi; \pi \in \mathcal{P}_{sym}(\Omega \times \Omega, dx) \right\}, \quad (36)$$

where the cost function is given by  $c(x, y) = \langle u(x), y \rangle$ . In the case where  $u$  is monotone then the involution where the supremum is attained is simply the identity.

### 3 Multidimensional Monge-Kantorovich Theorems

In this section, we are interested in relating the Gangbo-Świąch solution of the multidimensional Monge-Kantorovich with quadratic cost to the following recent result of Ghoussoub-Moameni [14].

**Theorem 3.1** *Given a probability measure  $\mu$  on  $\Omega$  and bounded vector fields  $u_1, u_2, \dots, u_{m-1}$  from  $\Omega$  to  $\mathbb{R}^d$  that are  $\mu$ -non-degenerate, then*

1. *The symmetric Monge-Kantorovich problem*

$$\mathbf{K}_{\text{sym}} = \sup \left\{ \int_{\Omega^m} [\langle u_1(x_0), x_1 \rangle + \dots + \langle u_{m-1}(x_0), x_{m-1} \rangle] d\pi; \pi \in \mathcal{P}_{\text{sym}}(\Omega^m, \mu) \right\} \quad (37)$$

*attains its maximum at a measure of the form  $\bar{\theta} = (I, S, S^2, \dots, S^{m-1})_{\#}\mu$ , where  $S$  is a  $\mu$ -measure preserving transformation on  $\Omega$  such that  $S^m = I$  a.e.*

2. *There exists a Hamiltonian  $H \in \mathcal{H}_m(\Omega)$  such that for  $i = 1, \dots, m-1$ ,*

$$u_i(x) = \nabla_{i+1} H(x, Sx, S^2x, \dots, S^{m-1}x). \quad (38)$$

*Moreover,  $H$  could be replaced by a Hamiltonian that is concave in the first variable and convex in the other variables, though only  $m$ -cyclically sub-antisymmetric.*

3. *If the vector fields  $u_1, u_2, \dots, u_{m-1}$  are  $m$ -cyclically monotone, then (38) holds with  $S$  being the identity.*

We note first that the general multi-marginal version of the Monge-Kantorovich problem **(K)** where the probability measures  $\mu_i$ ,  $i = 0, 1, \dots, m-1$  are given marginals on domains  $\Omega_i \subset \mathbb{R}^d$ , and where  $c$  is any bounded lower semi-continuous cost function  $c : \Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1} \rightarrow \mathbb{R}$ , has the following dual problem.

**Proposition 3.1** *There exists a solution  $\bar{\theta}$  to the Kantorovich problem **(K)**, as well as an  $m$ -tuple of functions  $(u_0, u_1, \dots, u_{m-1})$  such that for all  $i = 0, \dots, m-1$ ,*

$$u_i(x_i) = \inf_{\substack{x_j \in \Omega_j \\ j \neq i}} \left( c(x_0, x_1, \dots, x_{m-1}) - \sum_{j \neq i} u_j(x_j) \right), \quad (39)$$

*and which maximizes the following dual problem*

$$\sum_{i=0}^{m-1} \int_{\Omega_i} u_i(x_i) d\mu_i \quad (\mathbf{D})$$

*among all  $m$ -tuples  $(u_0, u_1, \dots, u_{m-1})$  of functions  $u_i \in L^1(\mu_i)$  for which  $\sum_{i=0}^{m-1} u_i(x_i) \leq c(x_0, \dots, x_{m-1})$  for all  $(x_0, \dots, x_{m-1}) \in \Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1}$ .*

*Furthermore, the maximum value in **(D)** coincides with the minimum value in **(K)**, and*

$$\sum_{i=0}^{m-1} u_i(x_i) = c(x_0, \dots, x_{m-1}) \quad \text{for all } (x_0, \dots, x_{m-1}) \in \text{support}(\bar{\theta}). \quad (40)$$

In their seminal paper, Gangbo and Świąch [7] dealt with the case of a quadratic cost function,

$$c(x_0, x_1, x_2, \dots, x_{m-1}) = \sum_{i=0}^{m-1} \sum_{j=i+1}^{m-1} |x_i - x_j|^2 \text{ on } \mathbb{R}^d \times \mathbb{R}^d \times \dots \times \mathbb{R}^d, \quad (41)$$

and established the following remarkable result.

**Theorem 3.2 (Gangbo-Świąch)** *Consider Borel probability measures  $\mu_i$  on domains  $\Omega_i \subset \mathbb{R}^d$ , for  $i = 0, 1, \dots, m-1$  vanishing on  $(d-1)$ -rectifiable sets and having finite second moments, and let  $c$  be a quadratic cost as in (41). Then,*

1. *There exists a unique measure  $\bar{\theta}$  on  $\Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1}$ , where **(K)** is achieved. It is of the form  $\bar{\theta} = (T_0, T_1, T_2, \dots, T_{m-1})_{\#}\mu_0$  on  $\Omega_0 \times \Omega_1 \times \dots \times \Omega_{m-1}$ , where  $T_0 = I$ ,  $T_i : \Omega_0 \rightarrow \Omega_i$  and  $T_{i\#}\mu_0 = \mu_i$ .*
2. *Each  $T_i$  is one-to-one  $\mu_i$ -almost everywhere, is uniquely determined, and has the form*

$$T_i(x) = \nabla f_i^*(\nabla f_1(x)) \text{ where } f_i(x) = |x|^2/2 + \varphi_i(x) \text{ for } x \in \mathbb{R}^d, \quad (42)$$

*and  $\varphi_i$  is a convex function which is related to the solutions  $(u_i)_{i=0}^{m-1}$  of the dual problem **(D)** by the formula  $\varphi_i(x) = \frac{m-1}{2}|x|^2 - u_i(x)$  for  $x \in \mathbb{R}^d$ .*

3. Moreover,  $\nabla f_0(x) = x + T_1x + T_2x + \dots + T_{m-1}x$  for  $\mu_0$ -almost all  $x \in \mathbb{R}^d$ .

The above result was clarified further by Agueh and Carlier [1], who essentially established the following.

**Proposition 3.1 (Agueh-Carlier)** *Under the conditions of the Theorem 3.2 and with the same notation, we have for each  $i = 0, \dots, m-1$ , that  $\frac{1}{m}\nabla f_i$  is the Brenier map that pushes the measure  $\mu_i$  onto the measure  $\nu$  on  $\mathbb{R}^d$  which is the image of the optimal measure  $\bar{\theta}$  by the “barycentric map”  $(x_0, \dots, x_{m-1}) \rightarrow \frac{1}{m} \sum_{i=0}^{m-1} x_i$ . Moreover,  $\nu$  is the unique minimizer of the functional  $\nu \rightarrow \sum_{i=0}^{m-1} W_2^2(\mu_i, \nu)$  where  $W_2$  is the Wasserstein distance.*

We now describe the situation in the case where the measures  $\mu_i$  are obtained from one measure  $\mu$  by cyclic permutations.

**Corollary 3.3** *Let  $\mu$  be a probability measure on  $\mathbb{R}^N$  and let  $\sigma$  be a unitary linear  $m$ -involution on  $\mathbb{R}^N$ , that is  $\sigma^* = \sigma^{-1}$  and  $\sigma^m(x) = x$ . Consider the corresponding Kantorovich problem associated to the measures  $\mu_i := \sigma_{\#}^i \mu$ ,  $i = 0, \dots, m-1$  with quadratic cost on  $\mathbb{R}^{mN}$ . Then,*

1. The barycentric measure  $\nu$  on  $\mathbb{R}^N$  associated to the measure  $\mu_i := \sigma_{\#}^i \mu$ ,  $i = 0, \dots, m-1$  is  $\sigma$ -invariant.
2. There exists a strictly convex function  $f_0 : \mathbb{R}^N \rightarrow \mathbb{R}$  such that the functions  $f_i(x) := f_0(\sigma^{m-i}(x))$  satisfy

$$\nabla f_{i\#} \mu_i = \nu \quad \text{for } i = 0, \dots, m-1. \quad (43)$$

**Proof:** As shown by Agueh-Carlier [1],  $\nu$  is the unique minimizer of the functional  $\nu \rightarrow \sum_{i=0}^{m-1} W_2^2(\mu_i, \nu)$ , where  $\mu_i := \sigma_{\#}^i \mu$ . Since  $\sigma^m = I$ , the uniqueness yields that  $\nu$  is then  $\sigma$ -invariant. Since now both functions  $f_i$  and  $\psi_i := f_0 \circ \sigma^{m-i}$  are convex and since both  $\nabla f_i$  and  $\nabla \psi_i = \sigma^i \circ \nabla f_0 \circ \sigma^{m-i}$  push  $\mu_i$  onto  $\nu$ , we get from the uniqueness property of Brenier maps that –modulo a constant–  $f_i = f_0 \circ \sigma^{m-i}$ .  $\square$

### $m$ -cyclically monotone operators

We shall now apply the above corollary to the case where  $\sigma$  is the cyclic permutation  $\sigma(x_0, \dots, x_{m-1}) = (x_1, \dots, x_{m-1}, x_0)$  on a product space  $X^m$ . But first, we point to the connection with  $m$ -cyclically monotone operators studied recently by Galichon-Ghoussoub [6]. They proved the following which is obviously an extension of Krauss’ theorem to the case when  $m \geq 3$ .

**Theorem 3.4 (Galichon-Ghoussoub)** *Let  $u_1, \dots, u_{m-1} : \Omega \rightarrow \mathbb{R}^d$  be  $m$ -cyclically monotone vector fields. Then, there exists a Hamiltonian  $H \in \mathcal{H}_m$  such that*

$$(u_1(x), \dots, u_{m-1}(x)) = \nabla_{x_2, \dots, x_m} H(x, x, \dots, x) \text{ for all } x \in \Omega. \quad (44)$$

Moreover,  $H$  can be replaced by a Hamiltonian  $K$  on  $\mathbb{R}^d \times (\mathbb{R}^d)^{m-1}$ , which is concave in the first variable, convex in the last  $(m-1)$  variables, whose restriction to  $\Omega^m$  is  $m$ -cyclically sub-antisymmetric. The concave-convex function  $K$  is  $m$ -cyclically antisymmetric in the following sense: For every  $\mathbf{x} = (x_0, \dots, x_{m-1})$  in  $\Omega^m$ , we have

$$K(x_0, x_1, \dots, x_{m-1}) + K_{2, \dots, m}(x_0, x_1, \dots, x_{m-1}) = 0 \quad (45)$$

where  $K_{2, \dots, m}$  is the concavification of the function  $L(\mathbf{x}) = \sum_{i=1}^{m-1} H(\sigma^i \mathbf{x})$  with respect to the last  $m-1$  variables.

As to the connection to measure preserving  $m$ -involutions, they also showed that  $u : \Omega \rightarrow \mathbb{R}^d$  is  $m$ -cyclically monotone  $\mu$  a.e. if and only if it is in the polar set of  $\mathcal{S}_m(\Omega, \mu)$ , that is  $\inf\{\int_{\Omega} \langle u(x), x - Sx \rangle d\mu; S \in \mathcal{S}_m(\Omega, \mu)\} = 0$ . Equivalently, the projection of  $u$  on  $\mathcal{S}_m(\Omega, \mu)$  is the identity map, i.e.,

$$\inf\left\{\int_{\Omega} |u(x) - Sx|^2 d\mu(x); S \in \mathcal{S}_m(\Omega, \mu)\right\} = \int_{\Omega} |u(x) - x|^2 d\mu(x).$$

It is easy to see that the above is also equivalent to the statement that

$$\sup\left\{\int_{\Omega^m}\langle u(x_0), x_{m-1}\rangle d\pi(\mathbf{x}); \pi \in \mathcal{P}_{\text{sym}}(\Omega^m, \mu)\right\} = \int_{\Omega}\langle u(x), x\rangle d\mu(x), \quad (46)$$

and that the sup is attained at the image of  $\mu$  by the map  $x \rightarrow (x, x, \dots, x)$ , which is nothing but a particular case of the symmetric Monge-Kantorovich problem, when the cost function is given by  $c(x_0, \dots, x_{m-1}) = \langle u(x_0), x_{m-1}\rangle$ , where  $u$  is an  $m$ -cyclically monotone operator.

## Multidimensional Monge theorem on graphs and $m$ -involutions

We shall now make a connection between Theorem 3.1 above, which is a Monge-Kantorovich problem on  $\mathcal{P}_{\text{sym}}(\Omega^m, \mu)$  with

$$c(x_0, x_1, \dots, x_{m-1}) = \langle u_1(x_0), x_1\rangle + \dots + \langle u_{m-1}(x_0), x_{m-1}\rangle$$

as a cost function, and the mass transport result of Gangbo-Świąch, which corresponds to the standard multidimensional Monge-Kantorovich problem, i.e., with cost function

$$c(y_0, y_1, \dots, y_{m-1}) = \sum_{i=0}^{m-1} \sum_{j=i+1}^{m-1} |y_i - y_j|^2,$$

and where the marginals are  $\mu_i := \sigma_{\#}^i \mu_0$  and  $\mu_0$  is supported on an appropriate graph dictated by the vector fields  $u_1, \dots, u_{m-1}$ .

For simplicity, we shall do this for  $m = 3$ , that is for two vector fields  $u_1, u_2$  in  $L^\infty(\Omega; \mathbb{R}^d)$ . Consider the (degenerate) probability measure  $\mu$  to be the image of Lebesgue measure  $dx$  on  $\Omega \subset \mathbb{R}^d$  by the map  $P : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^{2d} \times \mathbb{R}^{2d} \times \mathbb{R}^{2d}$  defined by

$$x \rightarrow P(x) = (x, x, u_1(x), 0, 0, u_2(x)).$$

We now consider the 3-cyclic permutation on  $\mathbb{R}^{2d} \times \mathbb{R}^{2d} \times \mathbb{R}^{2d}$  defined by

$$\sigma((x_{0,1}, x_{0,2}), (x_{1,1}, x_{1,2}), (x_{2,1}, x_{2,2})) = ((x_{1,1}, x_{1,2}), (x_{2,1}, x_{2,2}), (x_{0,1}, x_{0,2})),$$

in such a way that  $\sigma^3 = \text{Id}$ .

The quadratic 3-dimensional Monge problem of Gangbo-Świąch applied to the measures  $\mu_0 = \mu$ ,  $\mu_1 = \sigma_{\#} \mu$  and  $\mu_2 = \sigma_{\#}^2 \mu$  becomes the problem of minimizing

$$\int_{\Omega} \{ \|P(x) - \sigma P(S_1 x)\|^2 + \|\sigma P(S_1 x) - \sigma^2 P(S_2 x)\|^2 + \|\sigma^2 P(S_2 x) - P(x)\|^2 \} dx$$

over all measurable maps  $(S_1, S_2)$ , where  $S_i : \Omega \rightarrow \Omega$  is measure preserving for  $i = 1, 2$ .

The Kantorovich formulation of the problem is then to minimize

$$C(\pi) := \int_{(\mathbb{R}^d)^3} \{ \|P(x) - \sigma P(y)\|^2 + \|\sigma P(y) - \sigma^2 P(z)\|^2 + \|\sigma^2 P(z) - P(x)\|^2 \} d\pi(x, y, z)$$

over all probability measures  $d\pi(x, y, z)$  whose 3 marginals are Lebesgue measure. In other words,

$$\begin{aligned} C(\pi) = \int_{(\mathbb{R}^d)^3} \{ & |x - u_1(y)|^2 + |x|^2 + |u_1(x)|^2 + |u_2(y)|^2 + |y|^2 + |u_2(x) - y|^2 \\ & + |u_1(y)|^2 + |u_2(z)|^2 + |z|^2 + |u_2(y) - z|^2 + |y - u_1(z)|^2 + |y|^2 \\ & + |x|^2 + |x - u_2(z)|^2 + |u_1(x) - z|^2 + |z|^2 + |u_1(z)|^2 + |u_2(x)|^2 \} d\pi(x, y, z). \end{aligned}$$

Since the integrals of  $x^2, y^2, z^2, u_1(x)^2, u_1(y)^2, u_1(z)^2, u_2(x)^2, u_2(y)^2, u_2(z)^2$  against the given marginals of  $\pi$  are given constants, the above problem amounts to minimize

$$\int_{(\mathbb{R}^d)^3} \{ |x - u_1(y)|^2 + |u_2(x) - y|^2 + |u_2(y) - z|^2 + |y - u_1(z)|^2 + |x - u_2(z)|^2 + |u_1(x) - z|^2 \} d\pi(x, y, z),$$

or, for the same reasons, to maximize

$$D(\pi) := \int_{(\mathbb{R}^d)^3} \{ \langle u_1(y), x \rangle + \langle u_2(x), y \rangle + \langle u_2(y), z \rangle + \langle u_1(z), y \rangle + \langle u_2(z), x \rangle + \langle u_1(x), z \rangle \} d\pi(x, y, z),$$

which is exactly the problem  $(\mathbf{K}_{\text{sym}})$  where the cost

$$c(x, y, z) = \langle u_2(x), y \rangle + \langle u_1(x), z \rangle$$

has been symmetrized. Consider now the optimal maps  $T_1, T_2$  obtained by Gangbo-Święch, that is  $T_i = \nabla f_i^* \circ \nabla f_0$  pushes  $\mu_0$  to  $\mu_i = \sigma_{\#}^i \mu_0$  and where  $f_i$  is strictly convex for  $i = 1, 2$ . By Corollary 3.3, we have that  $f_0 = f_i \circ \sigma^i$  for  $i = 1, 2$ .

Note now that there exist measure preserving transformations  $S_1, S_2$  on  $\Omega$  such that the optimal maps

$$\tilde{T}_1 x := T_1(Px) = \nabla f_1^* \circ \nabla f_0(Px) = \sigma P(S_1 x) = (u_1(S_1 x), 0, 0, u_2(S_1 x), S_1 x, S_1 x)$$

maps  $dx$  onto  $\mu_1$ , while

$$\tilde{T}_2 x := T_2(Px) = \nabla f_2^* \circ \nabla f_0(Px) = \sigma^2 \circ P(S_2 x) = (0, u_2(S_2 x), S_2 x, S_2 x, u_1(S_2 x), 0)$$

maps  $dx$  onto  $\mu_2$ .

Let now  $T_{2,1} := \nabla f_2^* \circ \nabla f_1$  in such a way that  $\tilde{T}_2 = T_{2,1} \circ \tilde{T}_1$ , and let  $S_{2,1}$  be a measure preserving on  $\Omega$  such that

$$T_{2,1}(u_1(y), 0, 0, u_2(y), y, y) = (0, u_2(S_{2,1}y), S_{2,1}y, S_{2,1}y, u_1(S_{2,1}y), 0),$$

so that

$$T_{2,1} \circ \tilde{T}_1 x = (0, u_2(S_{2,1} \circ S_1 x), S_{2,1} \circ S_1 x, S_{2,1} \circ S_1 x, u_1(S_{2,1} \circ S_1 x), 0).$$

Since  $\tilde{T}_2 = T_{2,1} \circ \tilde{T}_1$ , we have that  $S_{2,1} \circ S_1 = S_2$ , and since  $\nabla f_0 = \sigma^{3-i} \circ \nabla f_i \circ \sigma^i$  for  $i = 1, 2$  and  $T_{2,1} \circ \sigma(Px) = \nabla f_2^* \circ \nabla f_1 \circ \sigma(Px)$  for a.e.  $x \in \Omega$ , one can easily verify that

$$S_{2,1} = S_1 := S, S_2 = S_{2,1} \circ S_1 = S^2 \text{ and } S^3 = I.$$

In other words, the points  $(t_0, T_1 t_0, T_2 t_0)$  are such that

$$t_0 = P(x) = (x, x, u_1(x), 0, 0, u_2(x)),$$

$$T_1 t_0 = \sigma P(Sx) = (u_1(Sx), 0, 0, u_2(Sx), Sx, Sx)$$

and

$$T_2 t_0 = \sigma^2 P(S^2 x) = (0, u_2(S^2 x), S^2 x, S^2 x, u_1(S^2 x), 0)$$

where  $S$  is a measure preserving transformation such that  $S^3 = I$ .

The convex function  $\Phi_0(t) = f_0(t) - \frac{1}{2}|t|^2$  is such that

$$\nabla \Phi_0(t_0) = T_1 t_0 + T_2 t_0 = (u_1(Sx), u_2(S^2 x), S^2 x, S^2 x + u_2(Sx), Sx + u_1(S^2 x), Sx).$$

Define now the convex Lagrangian

$$L(x, y, z) = \Phi_0(x, x, y, 0, 0, z) =: \Phi(t).$$

We then get

$$\nabla L(x, y, z) = (D_{0,1}\Phi(t) + D_{0,2}\Phi(t), D_{1,1}\Phi(t), D_{2,2}\Phi(t))$$

and

$$\nabla L(x, u_1(x), u_2(x)) = (u_1(Sx) + u_2(S^2 x), S^2 x, Sx).$$

Let now  $H$  be the Legendre transform of  $L$  with respect to the last two variables, that is

$$H(x, y, z) = \sup\{ \langle y, p \rangle + \langle z, q \rangle - L(x, p, q); p \in \mathbb{R}^d, q \in \mathbb{R}^d \}.$$

$L$  is clearly concave in the first variable, convex in the last two variables and

$$L(x, u_1(x), u_2(x)) + H(x, S^2 x, Sx) = \langle u_1(x), S^2 x \rangle + \langle u_2(x), Sx \rangle,$$

and in other words,  $(u_1(x), u_2(x)) = \nabla_{2,3} H(x, S^2 x, Sx)$  for all  $x \in \Omega$ .

## References

- [1] Agueh, M. and Carlier, G., *Barycenters in the Wasserstein space*, SIAM J. Math. Anal., 43(2) (2011) 904–924.
- [2] Burachik, R. S. and Svaiter, B. F., *Maximal monotonicity, conjugation and the duality product*, Proc. Amer. Math. Soc. **131** (2003), no.8, 2379–2383.
- [3] Brenier, Y. *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), 375–417.
- [4] Carlier, G. and Nazaret, B., *Optimal transportation for the determinant*, ESAIM Control Optim. Calc. Var., 14(4) (2008), 678–698
- [5] Fitzpatrick, S. P. *Representing monotone operators by convex functions*, Proc. Centre for Math. Analysis **20** (1989), 59–65.
- [6] Galichon, A. and Ghoussoub, N., *Variational representations for  $N$ -cyclically monotone vector fields*, Preprint (2012)
- [7] Gangbo, W. and Świąch, A., *Optimal maps for the multidimensional Monge-Kantorovich problem* Comm. Pure Appl. Math., 51(1) (1998) 23–45.
- [8] Ghoussoub, N., *Anti-selfdual Lagrangians: Variational resolutions of non self-adjoint equations and dissipative evolutions*, AIHP-Analyse non linéaire, 24 (2007), 171–205.
- [9] Ghoussoub, N., *A variational theory for monotone vector fields*, Journal of Fixed Point Theory and Applications, Volume 4, Number 1 (2008), 107–135
- [10] Ghoussoub, N., *Anti-symmetric Hamiltonians: Variational resolution of Navier-Stokes equations and other nonlinear evolutions*, Comm. Pure & Applied Math., vol. 60, no. 5 (2007), 619–653
- [11] Ghoussoub, N., *Selfdual partial differential systems and their variational principles*, Springer Monograph in Mathematics, Springer-Verlag (2008).
- [12] Millien, P., *A functional analytic approach to the selfdual polar decomposition*, Masters Thesis, UBC (2011).
- [13] Ghoussoub, N. and Moameni A., *A Self-dual Polar Factorization for Vector Fields*, Comm. Pure. Applied. Math (2012) 21pp (to appear)
- [14] Ghoussoub, N. and Moameni A., *Symmetric Monge-Kantorovich problems and polar decompositions of vector fields*, Preprint (2012)
- [15] Krauss, E., *A representation of arbitrary maximal monotone operators via subgradients of skew-symmetric saddle functions*, Nonlinear Anal. 9, no. 12 (1985), 1381–1399.
- [16] Pass, B., *Optimal transportation with infinitely many marginals*, Preprint (2012).
- [17] Pass, B., *PhD thesis*, University of Toronto, 2011.
- [18] Pass, B., *Uniqueness and Monge solutions in the multimarginal optimal transportation problem*, SIAM Journal on Mathematical Analysis, 43(6) (2011), 2758–2775
- [19] Pass, B., *On the local structure of optimal measures in the multi-marginal optimal transportation problem*, Calculus of Variations and Partial Differential Equations, 43:529–536, 2012. 10.1007/s00526-011-0421-z.
- [20] Phelps, R. R., *Convex functions, monotone operators and differentiability*, Lecture Notes in Math. 1364, Springer Verlag, New York, Berlin, Tokyo, (1998), 2nd edition 1993.

- [21] Rockafellar, T. *Convex Analysis*, Princeton University Press, 1970.
- [22] Svaiter, B. F., *Fixed points in the family of convex representations of a maximal monotone operator*, Proc. Amer. Math. Soc. **131**, no 12 (2003), 3851-3859.
- [23] Villani, C., *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2003.
- [24] Villani, C., *Optimal transport: old and new*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, New York, 2009.